

### III. REMARKS

Claims 1-13 are pending in this Application. By this Amendment, claims 1, 5, and 9 have been amended, and no claims have been canceled. Applicants are not conceding in this application that those claims are not patentable over the art cited by the Office, as the present claim amendments are only for facilitating expeditious prosecution of the subject matter. Applicants do not acquiesce in the correctness of the rejections and reserve the right to present specific arguments regarding any rejected claims not specifically addressed. Further, Applicants reserve the right to pursue the full scope of the subject matter of the original claims in a subsequent patent application that claims priority to the instant application. Reconsideration in view of the following remarks is respectfully requested.

#### **Rejections under 35 U.S.C. § 101**

In the Office Action, claims 1-4 are rejected under 35 U.S.C. § 101 as allegedly being directed to non-statutory subject matter. Specifically, the Office asserts that claim 1 neither transforms underlying subject matter nor positively recites structure associated with another statutory category (such as physical hardware).

With respect to independent claim 1, Applicants have amended the same herein to recite the features of “generating on a computer, for each of the orthographic variations, a cut and paste code” and “storing on a data carrier, for each of the orthographic variations, the associated extended cut and paste code, the data carrier including one of a magnetic computer disk and an optical computer disk.” (Claim 1.) No new subject matter is added in the course of these amendments, as support may be found in the specification as filed in at least paragraph [0052]. Applicants respectfully submit that these features more clearly tie the claimed computer implemented method to another statutory category, e.g., a manufacture or machine such as “a

computer” and “a data carrier including one of a magnetic computer disk and an optical computer disk,” and positively recites structure associated with another statutory category. Accordingly, Applicants submit that amended claim 1 recites patent-eligible subject matter.

Applicants respectfully submit that claims 2-4, which incorporate the above-described features of claim 1, also recite patent-eligible subject matter for the reasons stated above.

**Rejections under 35 U.S.C. § 103(a)**

In the Office Action, claims 1-13 are rejected under 35 U.S.C. § 103(a). Claims 1, 3, 5, 7, 9, 11, and 13 are rejected as being allegedly unpatentable over Kaplan *et al.* (US Pat. 5,594,641, hereinafter, “Kaplan”) in view of Holtz (*Data Compression for Disc Files and Communication Networks* (1993), hereinafter, “Holtz”); claims 2, 6, and 10 are rejected as being allegedly unpatentable over Kaplan in view of Holtz and Lee *et al.* (US Pat. 4,939,639, hereinafter, “Lee”); and claims 4, 8, and 12 are rejected as being allegedly unpatentable over Kaplan in view of Holtz and Schabes *et al.* (US Pat. 6,424,983, hereinafter, “Schabes”). Applicants have amended each of independent claims 1, 5, and 9 herein in order to improve clarity, and submit that the Kaplan and Holtz, either alone or in combination, do not teach or suggest each and every feature of the claimed invention.

For example, Applicants submit that the combination of Kaplan and Holtz does not teach, suggest, or make obvious the feature of “generating, for each or the orthographic variations, a cut and paste code, wherein the cut and paste code includes code which indicates a quantity of characters that should be cut from the end of a surface form of a word and pasted to produce a particular variation” (claim 1, and similarly recited in claims 5 and 9). In the Office Action, the Office relies on Kaplan to teach this feature.

Unlike the claimed invention, however, Kaplan's technique for text indexing and retrieval relies not on a "cut and paste code," among other features, but on sets of ordered pairs encoded by a finite state transducer (FST), such as {<arrive arrive>, <arrive arriving>, <arrive arrived>, <arrive arrives>, <arrive arrival>}. (Kaplan, Col. 4, lines 19-23.) Kaplan's FSTs represent any "regular relation" including finite lists of ordered string pairs, in addition to "certain kinds of infinite collections of ordered pairs." (*Id.*, lines 23-26.) The ordered pairs are encoded as an FST data structure, which can be used by a computer processor to get all the items associated with any given input. Using the example of ordered string pairs above, this means that given the stem "arrive," one could get all of the various forms ("arriving," "arrived," "arrives," "arrival," etc.); or given one of the variant forms ("arrival"), one could get its stem ("arrive"). (*Id.*, lines 29-32.)

In the Office Action, the Office asserts that Kaplan's tag codes "indicate operations which are to be performed on a base stem in order to arrive at a particular orthographic variation." (Office Action, p. 3.) Applicants submit that the tags taught by Kaplan are "morphological tags," such as "+PresPart" which "encode[s] the information that 'arriving' is the present participle form of 'arrive'" (col. 4, line 62-64). As such, Kaplan's tags do not indicate operations at all. Further, even if, *arguendo*, these tag codes did have the effect of, e.g., "indicat[ing] that an 'e' is to be removed or 'cut' (i.e., 1 letter) from the base form 'arrive' and then 3 letters, 'ing', are to be added to the modified base form" (Office Action, p. 3), this does not teach a cut and paste code including "code which indicates a quantity of characters that should be cut from the end of a surface form of a word and pasted to produce a particular variation" as claimed herein.

As taught by Kaplan, a two-level FST can be used to map a particular word to its lexical representation, as depicted in Kaplan's FIG. 2 (continuing with the "arrive" stem example

described above). In FIG. 2, the start state is labeled with an 's', and the end state is labeled with a '©' symbol. The 'ε' (epsilon) symbols, used in place of letters at the end of the shorter word ("arrive"), act as a null symbol to allow the FST to continue to process even where the stem word ("arrive") and the generated word ("arriving") do not have a comparable number of letters. (See, Kaplan, FIG. 2; col. 4, lines 40-44.) However, this insertion of an ε character allowing the FST to continue to process as described above, fails to teach "code which indicates a quantity of characters that should be cut from the end of a surface form of a word and pasted to produce a particular variation" as claimed herein, because there is no quantity of characters involved. Instead of indicating a quantity of characters to be cut or pasted to a word, the ε character may ostensibly be used as the final character following the cutting or pasting of any quantity of characters.

Kaplan further teaches that additional FSTs may be created to apply additional rules, as taught in Kaplan, col. 7-8:

FIG. 5 shows an example of an FST created to generate stems or lexical forms from a surface form based on a rule. ... A sample two-level rule in words: Lexical N is realized as surface m if and only if it is followed by a p on the lexical side (regardless of how that p itself is realized). ... This rule might be used to indicate that the words "impractical" and "intractable" are derived from the lexical forms "iNpractical" and iNtractable", respectively, which have the same lexical representation (iN) for the prefix meaning "not". The corresponding transducer shown in FIG. 5 comprises a state 50, the double circle that is both the start and a final state. If N is realized as m, the transducer moves to state 51, a nonfinal state which requires a p as the next input. If N is realized as n, the transducer moves to state 52 that does not allow a p to follow. The transducer thus encodes one additional fact that is not explicitly represented in the rule: lexical N's are realized as n in all other cases. (Col. 7, line 56 – Col. 8, line 9.)

Applicants submit that these teachings also fail to disclose the afore-described claimed feature, as the characters to be cut or pasted to a word are not indicated by "code which indicates a

quantity of characters that should be cut from the end of a surface form of a word and pasted to produce a particular variation,” but rather, by a character that precedes the particular character.

With further respect to claims 1, 5, and 9, Applicants submit that the proposed combination of Kaplan and Holtz fails to teach or suggest “generating ... a cut and paste code, ... wherein the cut and paste code is extended by a gloss code that indicates whether at least part of the orthographic variation should be converted between upper and lower case” (claim 1, and similarly recited in claims 5 and 9). In the Office Action, the Office relies on Holtz to teach this feature.

Holtz teaches compression of data through use of a string compression system (p. 146). In one embodiment, a library is generated off-line from random text files. Each unique text word will generate a 16 bit output code which will cause a list search to find that code already stored. (p. 148.) Relative to the case of characters, Holtz teaches that “for a more compact library which contains only lower case words two extra bits are added. The FC bit will indicate that the first character of a word is a capital character. The AC bit indicates that that the whole word consists of capital letters. If both the AC and the FC bit is set then a special code is send [*sic*] to the receiver.” (p. 149, right col.)

Applicants submit, however, that neither this passage nor the balance of Holtz teaches a cut and paste code “extended by a gloss code that indicates whether at least part of the orthographic variation should be converted between upper and lower case.” Holtz merely teaches an extra two bits of code per word to be stored in a library (p. 148). In view of this, and in view of Kaplan's failure to teach a copy and paste code, Applicants respectfully submit that Kaplan and Holtz fail to teach or suggest “generating ... a cut and paste code ... wherein the cut and paste code is extended by a gloss code that indicates whether at least part of the orthographic

variation should be converted between upper and lower case.” Still further, Applicants submit that it follows from the above that Kaplan and Holtz also do not teach or suggest “storing ... for each of the orthographic variations, the associated extended cut and paste code...”

In view of at least the amendments contained herein and the remarks above, Applicants respectfully submit that the rejections of claims 1, 5, and 9 under § 103(a) are defective, and request withdrawal of the same.

With regard to dependent claims 2-4, 6-8, and 10-13, Applicants respectfully submit that these claims are allowable for reasons stated above relative to independent claims 1, 5, and 9, as well as for their own additional claimed subject matter. With further regard to dependent claims 2, 6, and 10; and 4, 8, and 12; Applicants additionally submit that Lee and Schabes respectively fail to cure the defects described above with regard to claims 1, 5, and 9. Accordingly, Applicants respectfully request that the Office withdraw the rejections under 35 U.S.C. § 103(a) to claims 2-4, 6-8, and 10-13.

#### IV. CONCLUSION

Applicants respectfully submit that the Application as presented is in condition for allowance. Should the Examiner believe that anything further is necessary in order to place the application in better condition for allowance, the Examiner is requested to contact Applicants' undersigned attorney at the telephone number listed below.

Respectfully submitted,

/Jayme M. Torelli/

Jayme M. Torelli  
Reg. No.: 62,735

Dated: June 3, 2009

Hoffman Warnick LLC  
75 State Street, 14<sup>th</sup> Floor  
Albany, NY 12207  
Phone: (518) 449-0044  
Fax: (518) 449-0047